ESG and earnings quality: pillar decomposition and greenwashing stability

João Fragoso
Universidade Europeia, Lisbon, Portugal

Carlos Manuel Pinheiro
Universidade Europeia, Lisbon, Portugal | Instituto Universitário de Lisboa (ISCTE-IUL) Business Research Unit (BRU-IUL), Lisbon, Portugal | ISCAL, Universidade Politécnica de Lisboa, Portugal | CETRAD, Vila Real, Portugal |

Abstract

Purpose – This paper investigates whether the relationship between environmental, social, and governance (ESG) performance and earnings quality varies depending on the ESG pillar, earnings-quality attribute or temporal horizon, or is symbolically uniform and compatible with a population of firms that engage in greenwashing.

Design/methodology/approach – The study draws on Eurozone listed firms covered by Refinitiv (LSEG) over 2004–2023. Earnings quality is measured using five complementary attributes: discretionary accruals, accrual quality, persistence, predictability, and smoothness. Pillar-level relationships are estimated with Fama-MacBeth cross-sectional regressions and triangulated with two-way fixed-effects, quantile, and matching estimators (propensity score matching and entropy balancing), with difference-in-differences as identification support and an instrumental-variables check reported in full in the online appendix. Greenwashing is defined as high declared ESG performance paired with high discretionary accruals against year-specific medians. A fixed-threshold classification of the same firm-years provides a contrasting series.

Findings – The relationship is structured rather than uniform. The Social pillar is positively and significantly associated with earnings persistence, the Governance pillar is the only pillar significantly associated with reduced masking and scaled net income volatility; and the Environmental pillar is associated with improved accrual quality. This structure coexists with a stable core. Greenwashing prevalence, measured against contemporaneous benchmarks, remains between 21.0 and 24.1 per cent of firm-years over 2008–2023. There was no structural break when the Non-Financial Reporting Directive (NFRD, 2017) or the Corporate Sustainability Reporting Directive (CSRD, 2022) came into force, and it persists at the firm level. When a fixed threshold is applied, the measured share rises from 16.3 to 25.9 per cent. The stability of the year-specific series shows that measured prevalence is robust to the long-term rise in ESG scores; the divergence between the two series isolates the distortion that this rise induces under relative scoring.

Research limitations/implications – The study relies on ESG scores made available by a single rating provider with a relative scoring methodology, covers listed firms in the Eurozone, and its design is associative rather than causal. Future research could replicate the classification with alternative ESG data providers, port the year-specific versus fixed-threshold contrast to other relative scoring systems, and revisit the invariance as the post-CSRD window lengthens.

Practical implications – Disclosure mandates appear to reshape reporting practice without dislodging a persistent core of greenwashing firms, which implies that mandates move average reporting levels more readily than configurations. Investors and analysts should benchmark ESG scores against the contemporaneous distribution, since a fixed benchmark overstates the growth of greenwashing, and should read pillar scores as signals for distinct earnings-quality attributes.

Originality/value – To the best of our knowledge, this is the first study to jointly map the three pillars of ESG (environment, social, governance) onto five attributes of earnings quality across two time horizons in a European context. The study also tests for greenwashing in a way that is robust to the long-term increase in ESG scores. The identification strategy contrasts classifications based on annual medians with a fixed pooled threshold; this contrast is offered as a measurement contribution for research that classifies firms under relative scoring systems.

Keywords: ESG ratings; earnings quality; greenwashing; pillar decomposition; disclosure regulation; Eurozone

JEL classification: M41; M14; G32; G38; Q56

1. Introduction

Whether firms with strong ESG credentials also deliver disciplined financial reporting has become an urgent question due to regulation. Disclosure requirements now cover most listed companies in Europe. Mandatory environmental, social, and governance (ESG) disclosure improves stock liquidity across a panel of sixty-five countries, with the strongest effects where mandates are applied on a full-compliance rather than comply-or-explain basis (Krueger et al., 2024). However, does ESG disclosure go along with credible financial statements? This is an unsettled issue still requiring analysis.

Two bodies of research on the relationship between ESG performance and earnings quality disagree about structure. Throughout the paper, ESG performance refers to the ESG scores assigned to the firm by a rating provider on the basis of what the firm reports, an external assessment of its sustainability profile rather than the disclosure itself. According to the substantive view (Hong & Andersen, 2011; Kim et al., 2012), sustainability engagement is expected to discipline reporting through reputational capital and governance oversight; the relationship should then be structured, pillar by pillar, at the horizons where each source of discipline operates. In contrast, under the symbolic view (Suchman, 1995; Prior et al., 2008), engagement is merely for show, so the relationship should be uniform and weak, compatible with a stable sub-population of firms whose ESG scores exceed the quality of their reporting. Regulation may change how these firms are classified without reducing the size of the group.

Recent reviews record how little of the extant evidence engages either view at the required granularity: the literature has predominantly relied on composite scores, with limited attention to the three pillars, and the decoupling between CSR rhetoric and actual practices remains underexplored (Deng et al., 2024). The most recent synthesis of this decoupling calls for measures that capture the persistence of the gap between disclosure and performance rather

than its mere presence (Ștefănescu & Suciu, 2026); the classified, year-specific prevalence examined here answers that call.

The core question addressed in this paper is whether the relationship between ESG performance and earnings quality is substantively structured and specific to a given pillar, earnings-quality dimension, and temporal horizon, or symbolically uniform, a reputational "greenwash" that coexists with a stable population of greenwashers across regulatory regimes.

Using Refinitiv (London Stock Exchange Group, LSEG) data on listed firms in the nineteen Eurozone countries over 2004–2023, the paper measures earnings quality as a vector of five attributes, decomposes the composite ESG score into its three pillars, and classifies as greenwashing the firm-years that pair a high declared ESG score with high discretionary accruals, both assessed against year-specific medians. On the first margin the answer is structured: the Social pillar carries persistence with the highest precision in the matrix, the Governance pillar alone is significantly associated with the volatility of scaled net income, in the direction reduced masking predicts, and the Environmental pillar alone is significantly associated with accrual quality, in the quality-improving direction. On the second margin the answer is symbolic: the greenwasher share stays between 21 and 24 per cent across 2008–2023, crosses the entry into force of the Non-Financial Reporting Directive (NFRD) and the Corporate Sustainability Reporting Directive (CSRD) without a break, and is persistent at firm level. Under a fixed pooled threshold, the same share appears to expand; the divergence is the mark of long-term score inflation under relative scoring, not of changed conduct.

The stability result stands in deliberate contrast with Aboud et al. (2024), who report that Directive 2014/95 reduced ESG decoupling among covered firms. The tension is instructive rather than irreconcilable: the two results measure different objects, a continuous decoupling score against the prevalence of a classified configuration; the concluding section returns to

what each can show. On the structural margin the nearest antecedent is Palas et al. (2025), whose single-market, single-vendor, two-attribute design the map drawn here extends to three pillars, five attributes, and two temporal horizons in the Eurozone.

The evidence carries direct relevance for investors and analysts. Markets struggle to price greenwashing on their own: European stock prices do not react to media greenwashing announcements, and ESG ratings barely differentiate the firms exposed (Teti et al., 2024). Verification is what works: institutional investors who visit firms in person curb greenwashing among the firms they monitor (Chen et al., 2026). A classification separating persistent greenwashers from disciplined reporters on audited accounting data extends that verification logic to outsiders, supplying information that prices do not impound. The pillar map adds a second filter: each pillar is associated with a distinct earnings-quality attribute at a distinct horizon, so pillar scores carry attribute-specific information that the composite conceals.

The paper makes three contributions. To the best of our knowledge, it offers the first joint map of the ESG relationship onto pillar, earnings-quality attribute, and temporal horizon, converting the substantive-versus-symbolic tension into an observable pattern. It treats the temporal stability of greenwashing prevalence as a formal hypothesis, tested against the two European disclosure regimes. And it introduces the fixed-versus-year-specific contrast as an identification device for the measurement distortion that relative scoring generates under long-term score inflation. A related study by Fragoso & Pinheiro (2026) reports evidence that regulation compresses the disciplinary margin of the ESG relationship; the evidence here indicates that the same pressure leaves the symbolic core in place: regulation reshapes the surface of sustainability reporting without removing its substrate.

The remainder of the paper is organised as follows. Section 2 derives the hypotheses from the substantive and symbolic views. Section 3 describes the sample, the measures, and the

estimation strategy. Section 4 reports the results and the robustness battery. Section 5 concludes.

2. Literature review and hypotheses development

2.1 ESG and earnings quality: substantive discipline versus symbolic conformity

Two competing views organise the existing research on the relationship between ESG performance and earnings quality. The distinction between substantive and symbolic legitimation is classic and still being refined (Ashforth & Gibbs, 1990; Gu & Matisoff, 2026). Under the substantive view, sustainability engagement involves binding commitments to stakeholders whose demands carry power, legitimacy, and urgency (Donaldson & Preston, 1995; Mitchell et al., 1997), and the credibility of these commitments is tied to reporting integrity: a firm that invests in stakeholder relationships stands to lose their accumulated value if caught misreporting. In the signalling variant, ESG performance operates as a costly signal that narrows information asymmetry (Connelly et al., 2011), a signal that unreliable financial statements diminish.

Two mechanisms are held to enforce the resulting discipline. The reputational mechanism rewards honesty: reliable conduct accrues quality-assuring premia, while firms caught misreporting bear market losses far exceeding the direct legal penalties (Klein & Leffler, 1981; Karpoff et al., 2008). The governance mechanism is one of oversight: sustainability engagement tends to be accompanied by stronger board and audit-committee monitoring, structures associated with lower abnormal accruals and closer oversight of the reporting process (Jensen & Meckling, 1976; Klein, 2002; Carcello et al., 2006). Board capital matters here: among FTSE-350 firms, nonexecutive directors attentive to ESG issues are associated with stronger corporate ESG performance (Clare et al., 2025).

Support for the disciplinary view comes from United States studies based on Kinder, Lydenberg, and Domini (KLD) ratings. Hong and Andersen (2011) report that socially responsible firms exhibit lower accrual-based earnings management and higher accrual quality, and Kim et al. (2012) document the same configuration over 18,160 firm-year observations between 1991 and 2009, together with fewer securities-enforcement actions. The international evidence is less consistent: across 1,653 firms in forty-six countries, corporate social responsibility mitigates smoothing and loss avoidance yet coexists with greater earnings aggressiveness, contingent on investor protection (Chih et al., 2008). Heterogeneity across measures within a single study suggests that the relationship is unlikely to be uniform across dimensions of earnings quality; the first hypothesis formalises this suggestion. Recent estimates reproduce the pattern in China and in the energy and utilities sector (Liu et al., 2025; Persakis et al., 2026).

The symbolic view is different. Legitimacy is a perception that can be managed through conformity of appearance rather than of substance (Suchman, 1995), and Prior et al. (2008) document a positive effect of earnings management on subsequent corporate social responsibility across 593 firms in twenty-six countries, interpreting this as managerial entrenchment: managers who manipulate earnings use sustainability engagement as a shield against stakeholder scrutiny, so the declared commitment ensures the misreporting instead of constraining it. Under this view, the relationship is superficial at most and fully compatible with a persistent sub-population of firms whose declared sustainability and reported quality point in opposite directions.

Regulation raises the stakes of the distinction. Mandatory ESG disclosure has moved from the periphery to the centre of European reporting policy (Desai & Vyas, 2026), and mandates carry real consequences: across staggered adoptions in sixty-five countries, stock liquidity improves after mandatory ESG disclosure, moderated by design and enforcement (Krueger et al., 2024).

However, whether regulatory pressure also dislodges symbolic conformity, rather than merely repricing disclosure, is a question the existing evidence leaves open. The remainder of this section converts the two views into testable hypotheses.

2.2 Earnings quality as a multidimensional vector

Earnings quality is not a unitary construct. Francis et al. (2004) characterise it as a vector of attributes, spanning accrual quality, persistence, predictability, and smoothness alongside market-based properties. Dechow et al. (2010) review the proxies and conclude that no single measure dominates, the appropriate choice being contingent on the decision context. Section 3.2 sets out the five attributes retained here.

A multidimensional dependent variable converts the tension between substance and symbolism into a testable prediction. If the relationship reflects symbolic conformity, it should be uniformly and weakly associated with the attributes, because nothing in a decorative commitment selects among the properties of the earnings process. Substantive engagement, in contrast, should be differentially associated through each mechanism, each leaving its imprint on the attribute closest to its object of discipline. The heterogeneity reported by Chih et al. (2008) is an early indication of precisely this pattern. A differential pattern across attributes, rather than the sign or magnitude of any single coefficient, is what distinguishes the two views.

The temporal horizon provides the second dimension of structure. Persistence is an accumulating property: substantive engagement becomes embedded in organisational processes and coupled with a longer planning horizon (Eccles et al., 2014), so its relationship with persistence should surface when comparing firms, the cumulative product of multi-year commitments rather than a year-on-year adjustment. Smoothness materialises differently, produced within the firm through contemporaneous accrual choices. Within-firm and cross-sectional estimates therefore answer different questions rather than approximating a common

parameter. Section 3.5 assigns each hypothesis to the estimator whose identifying variation matches its horizon; the hypotheses themselves follow.

2.3 The pillar decomposition

Aggregate ESG scores conceal as much as they reveal. Khan et al. (2016) document that the market prices financially material sustainability issues only, so a composite blends signals of different economic content; Berg et al. (2022) show that rating divergence across providers traces to scope, measurement, and aggregation weights, and varies across the three dimensions. Evidence of this kind removes any presumption that a composite coefficient estimates a homogeneous relationship.

Pillar-level evidence on the earnings-quality counterpart of ESG performance is, at best, scant. The nearest design, Palas et al. (2025), reports that the relationship among United States firms rated by MSCI concentrates in the Social pillar, measured through persistence and cash-flow predictability; the corroboration is welcome, and the mapping pursued here crosses three pillars with five attributes over two horizons, a joint decomposition for which, to the best of our knowledge, the extant literature offers no published counterpart.

The differentiated pattern anticipated in Section 2.2 is formalised as one umbrella hypothesis and predictions specific to each pillar.

H1. The cross-sectional relationship between ESG performance and earnings quality is pillar-specific: the three pillars are associated with distinct earnings-quality attributes rather than uniformly across all three.

The Social pillar aggregates the firm's commitments to employees, communities, customers, and human rights, the relationships in which the substantive view locates the value at risk from misreporting. Such commitments are cumulative, built over multi-year horizons, so their

imprint should surface in the durability of earnings across firms rather than in year-by-year accrual choices; the predicted absence of a smoothness relationship is part of the hypothesis, not a concession.

H1a. In the cross-section, the Social pillar is positively associated with earnings persistence, but not with earnings smoothness.

The Governance pillar proxies the strength of monitoring over the reporting process, from board structure to audit oversight. Where monitoring is weak, insiders smooth reported income to mask firm performance from outsiders (Leuz et al., 2003), and opacity of this kind permits the accumulation of withheld bad news (Jin & Myers, 2006); stronger governance should therefore travel with reported income that tracks underlying performance, which across firms translates into more volatile scaled net income rather than less.

H1b. In the cross-section, the Governance pillar is positively associated with the volatility of scaled net income, and it is the only pillar for which this relationship is statistically significant, consistent with reduced masking of underlying earnings volatility among well-governed firms.

One clause of precision is non-negotiable in what follows. The within-firm component of smoothing, identified by the within-firm estimators of Section 3.5, belongs to the reputational mechanism; the Governance pillar anchors the cross-sectional component of the smoothness relationship. The two components answer different questions at different horizons, so H1b is a claim about the cross-section alone, neither confirmed nor contradicted by within-firm estimates. The Environmental pillar completes the matrix without a lettered prediction: extant guidance is thin, so any relationship of the pillar is read as part of the differentiated pattern that H1 anticipates. Hypothesis 2 turns to the population that the symbolic view predicts.

2.4 Greenwashing: a relative concept in shifting distributions

Greenwashing refers to the discrepancy between a company's declared sustainability and its actual performance, measured at firm level as 'walk-versus-talk' (Huang et al., 2025). Yu et al. (2020) provide the seminal operationalisation, a peer-relative gap between a firm's disclosure and its performance; Zatini et al. (2026) generalise the logic into a talk-versus-walk matrix in which greenwashing risk sits where talk exceeds action; and text-based devices comparing the tone of CSR reports with realised performance extend the same family (Gao & Yin, 2025). The typology of Section 3.4 compares declared ESG performance with the quality of reported earnings; the configuration of interest pairs a high declaration with aggressive accruals.

The construct is relative at its core, and the reference distribution is itself changing. A firm greenwashes by presenting itself as better than the current standard, so a high declaration only has meaning against the contemporaneous distribution. Refinitiv ESG scores are ordinal and relative by design, produced through percentile ranking within peer groups whose composition changes over time, so the methodology shapes the interpretation of any threshold, cross-sectionally and over time (Benuzzi et al., 2025); rating disagreement widens precisely where firms disclose more, marking the scores as methodology-specific measures (Christensen et al., 2022). The rhetoric and the measurement interact in both directions: positive report tone widens disagreement across agencies (Luo et al., 2026), and rating divergence in turn intensifies the manipulation of narrative tone (Chen & Wang, 2025). In the European large-cap universe, apparent environmental performance rose by around forty per cent over 2015–2023 while the real component remained flat (Kathan et al., 2025). The measurement literature has accordingly converged on year-specific relative benchmarks, with industry-year medians the standard choice (Lublóy et al., 2025); a fixed threshold applied to a changing distribution alters its meaning, one cohort at a time.

Whether regulation dislodges this population is a contested issue, and the most compelling published evidence contradicts the symbolic account. Aboud et al. (2024) report that the mandatory regime of Directive 2014/95 reduced ESG decoupling among covered European firms, in a difference-in-differences design against US firms whose reporting remained voluntary; this paper challenges that result directly, and the reconciliation, which turns on the construct and the metric, is developed in the concluding section. Fiechter et al. (2022) document that the same directive increased the level of CSR activities, which leaves open whether configurations move as levels do. Evidence from the CSRD era keeps the question live: disclosure expansion concentrates in positive, non-sensitive indicators, and European circular economy reporting still asks whether disclosure signals substance or symbol (Leggerini et al., 2026; Zharfpeykan & Harrison, 2026).

The first greenwashing hypothesis is as follows: prevalence is measured as the proportion of firm-years that report high ESG performance alongside high accrual-based earnings management, compared to the year-specific medians in Section 3.4.

H2a. Greenwashing prevalence, classified against year-specific medians, remains stable between 2008 and 2023, with no significant change observed at the time of NFRD's (2017) or CSRD's (2022) implementation.

A design element accompanies the hypothesis and should not be read as one. The two classifications are computed side by side, pooled and year-specific, and the divergence between the resulting series carries the identifying content: under long-term score inflation the fixed-threshold series should drift upward while the contemporaneous-benchmark series stays flat, and the gap isolates the measurement distortion from any change in conduct. The year-specific benchmark is inherited from prior practice; the contribution lies in the contrast itself. The second leg of the symbolic account concerns persistence at firm level.

2.5 Greenwashing as a firm-level trait

A stable share is compatible with two very different populations, one churning and one persistent, and the symbolic account selects the second. If the pairing reflected transient circumstances, firms would rotate through the group and a flat aggregate would conceal high turnover; if it reflects a durable configuration of incentives and oversight, the same firms should reoccupy the group year after year. The available firm-level evidence favours persistence: a one-year lagged greenwashing score is among the strongest predictors of future greenwashing among European large caps (Eckberg et al., 2026); because that study shares authors, universe, and case base with Kathan et al. (2025), and measures hand-collected accusations rather than an accounting configuration, the two are read as a single evidentiary strand.

The second greenwashing hypothesis suggests that this persistence is a characteristic of the firm.

H2b. Greenwashing is a persistent trait at firm level: given the disclosure regime, the probability that a firm will be classified as a greenwasher in year t increases if it was classified as such in year t-1.

The conditioning clause is important: persistence is a claim about the firm rather than about the rulebook, and holding the disclosure regime fixed separates the trait from any level shift that regulation might induce. The tests of both greenwashing hypotheses are pre-specified in Section 3.5.

3. Data and method

3.1 Sample

The sample covers listed firms from the nineteen Eurozone countries over 2004–2023, sourced from Refinitiv (LSEG), which provides financial-statement data and ESG scores. The universe

comprises 42,685 firm-year observations. Requiring at least three consecutive firm-year observations, as the rolling-window attributes demand, removes 2,689 observations and leaves a full estimation panel of 39,996 firm-years across 2,837 firms, over which the industry-year accrual models are estimated. Restricting to firm-years with a Refinitiv ESG score removes a further 31,541 observations, yielding an ESG sub-panel of 8,455 observations on 901 firms; requiring the modified Jones accrual measure removes an additional 1,272, yielding the typology universe of 7,183 firm-years. The estimation conditions of the multivariate specifications then produce the final sample of 6,550 firm-years across 840 firms. This attrition is dominated by the lagged ESG regressor: 618 of the 633 firm-years lost lack the one-year lagged score, predominantly each firm's first year of ESG coverage, 15 lack revenue growth, and none fail the remaining criteria. The effective estimation window opens in 2006.

Three features of the design are consistent across specifications. First, the estimation samples are attribute-specific: the data requirements of the accrual, persistence, predictability, and smoothness attributes produce samples ranging from 4,775 to 7,070 firm-years, the typology universe holds 7,183, and each table reports its own N. Second, Germany, France and Italy jointly account for 60.4 per cent of the ESG sub-panel, mirroring the relative size of Eurozone equity markets. Third, financial firms are retained and identified by an indicator variable. Excluding them leaves the Governance and Environmental cells of the pillar map, the within-firm estimates, and the greenwashing typology materially unchanged, while the cross-sectional Social persistence coefficient attenuates; the sensitivity is reported in the Online Appendix (Table A2).

3.2 Earnings-quality attributes

Earnings quality is measured as a vector of five attributes, in line with the multidimensional strand of the measurement literature (Francis et al., 2004; Dechow et al., 2010). Accruals-based

earnings management (AEM_MJones) is the absolute value of discretionary accruals from the Modified Jones model (Jones, 1991; Dechow et al., 1995), estimated by ordinary least squares (OLS) within each industry-year; higher values indicate lower quality. Accrual quality (AQ_DD) is the absolute residual from the Dechow and Dichev (2002) mapping of working-capital accruals into past, current, and future operating cash flows; higher residuals indicate lower quality. Persistence (PER_phi) is the slope from a firm-level AR(1) regression of net income scaled by lagged total assets, estimated over rolling ten-year windows with a minimum of six observations. Predictability (PRED_sigma) is the standard deviation of the residuals from the same autoregression, lower values indicating more precise forecasts. Smoothness (SMO_NI) is the standard deviation of scaled net income, a broad-coverage variant that requires no cash-flow data; higher values indicate more volatile, less smoothed earnings, so the attribute is read jointly with the accrual measures rather than as a quality direction of its own.

3.3 ESG measurement

ESG performance is the Refinitiv ESG Score, a composite on a 0–100 scale, together with its three pillar scores; the principal regressor is the one-period lag of each score, mitigating contemporaneous reverse causality. The ESG Score is preferred to the ESG Combined Score, which penalises the headline score for media controversies, because the controversies signal enters separately through the typology of Section 3.4; folding it into the regressor would conflate declaration with verification. Reliance on a single provider is a deliberate design choice rather than an oversight: ESG ratings diverge across providers, and the divergence widens precisely where firms disclose more (Berg et al., 2022; Christensen et al., 2022), so holding the provider fixed keeps the classification margins internally consistent, and the cost of a methodology-specific score is addressed by the year-specific benchmarks of Section 3.4.

3.4 The greenwashing typology and the pooled/annual contrast

The typology classifies each firm-year on two declared-versus-delivered margins. The first tier crosses the ESG Score against the Modified Jones measure: a firm-year combining an ESG Score above the median with discretionary accruals above the median is classified as a greenwasher, the group in which a high declaration coexists with aggressive accounting. The refined tier adds the Controversies score and isolates, within the greenwasher group, the firm-years below the pooled controversies threshold, the sub-population for which an external verification signal contradicts the declaration.

The classification benchmarks are year-specific medians, a choice justified before any estimation on the conceptual grounds of Section 2.4: greenwashing is a relative construct, Refinitiv scores are themselves relative by construction, the product of a double percentile ranking within industry and country peer groups (Benuzzi et al., 2025), and a fixed historical threshold mechanically reclassifies later observations as the distribution rises (Berg et al., 2022; Kathan et al., 2025). Year-specific relative cuts, and industry-year medians in particular, are established practice (Lublóy et al., 2025). The controversies threshold, in contrast, is deliberately pooled: the distribution is degenerate, with its mass at 100, the value assigned to firm-years without recorded controversies, and a year-specific median would be unstable without adding information.

The pooled/annual contrast is retained as an identification device rather than as a superseded alternative: both classifications are computed in a single derived script, and the divergence between the two series, the contribution this paper claims, identifies the measurement distortion that long-term score inflation induces under a relative scoring methodology.

Hypothesis 2b requires a panel of classification transitions. Transitions are defined over consecutive firm-year pairs (a one-year gap) within 2008–2023, the window over which the

annual cross-sections are populated; the resulting transition panel comprises 5,989 pairs on 830 firms.

3.5 Estimation strategy

The estimation strategy allocates each hypothesis to the estimator whose identifying variation matches its horizon. The two-way fixed-effects estimator, with firm and year effects and firm-clustered standard errors, identifies within-firm covariation and serves as the within-horizon benchmark. The Fama and MacBeth (1973) procedure estimates a separate cross-sectional regression for each year, with a minimum of thirty firms per year, and aggregates the annual slopes through their cross-year average; it serves as the between-horizon benchmark and carries the pillar decomposition of H1.

The Canay (2011) two-step panel quantile estimator, with bootstrapped standard errors (500 replications), extends the within-firm evidence across the conditional distribution. Propensity score matching (Rosenbaum & Rubin, 1983) and entropy balancing (Hainmueller, 2012) complete the cross-sectional evidence: treatment is coverage by the Refinitiv ESG universe, matched one-to-one on size, leverage, profitability, revenue growth, the loss indicator, and the financial-sector indicator within a 0.20 standard-deviation caliper without replacement; 8,211 of the 8,455 ESG-covered firm-years stand against 28,103 candidate controls, yielding 3,959 matched pairs, and entropy balancing reweights the controls to match the first three moments of the treated distribution.

Control variables follow the earnings-quality convention (Francis et al., 2004; Dechow et al., 2010): firm size (log of total assets), leverage, return on assets, revenue growth, a loss indicator, and asset turnover, with market-to-book added where market data condition the specification. Continuous controls are winsorised at the first and ninety-ninth percentiles. A post-double-

selection LASSO exercise (Belloni et al., 2014) screens 190 numerical candidates after the within transformation and selects none, confirming the adequacy of the manual control set.

The tests of Hypothesis 2 are pre-specified. For H2a, the annual greenwasher share is regressed on a linear trend and regime indicators for the NFRD (2017 onwards) and the CSRD (2022 onwards), with Newey-West standard errors at lag two (Newey & West, 1987) and a weighted variant using annual sample sizes; Chow tests (Chow, 1960) are run at the two known dates, in the standard form at 2017 and in the predictive form at 2022, where the post-window spans only two years; Wilson 95 per cent confidence intervals are reported alongside, because non-rejection alone does not establish stability, and the Bai-Perron search (Bai & Perron, 1998) is supplementary given its low power at T = 16. For H2b, the year-on-year transition matrix is the principal result, accompanied by a pooled logit of status on lagged status and regime indicators with firm-clustered standard errors, the pooled form chosen because the estimand is total persistence and a fixed-effects logit would suffer incidental-parameters bias at short T; a fixed-effects linear probability model is reported as robustness. The register throughout is associative: no estimator is asked to carry a full causal claim, and divergences across estimators are read first as differences in estimands.

3.6 Descriptive statistics

Table 1 presents the descriptive statistics: Panel A covers the estimation sample of each attribute, winsorised as used in estimation; Panel B the ESG sub-panel before winsorisation; and Panel C the final sample of Section 3.1 (N = 6,550), winsorised as used in estimation.

Table 1. Descriptive statistics

Variable	N	Mean	S.D.
Panel A: Earnings-quality attributes (estimation samples)			
|DA| Modified Jones (AEM_MJones)	6,550	0.0583	0.0711
AQ Dechow-Dichev (AQ_DD)	4,775	0.0550	0.0601

Variable	N	Mean	S.D.
Persistence (PER_phi)	6,112	0.1381	0.5773
Predictability (PRED_sigma)	6,112	0.0301	0.0419
Smoothness (SMO_NI)	7,070	0.0353	0.0546
Panel B: ESG scores (ESG sub-panel)			
ESG Score	8,455	54.99	20.93
Environmental pillar	8,454	51.67	27.56
Social pillar	8,454	57.28	24.47
Governance pillar	8,455	48.73	23.07
Controversies score	8,454	91.24	21.99
Panel C: Controls (final sample, N = 6,550)			
Firm size (ln assets)	6,550	8.4494	1.6449
Leverage	6,550	0.2845	0.1664
Return on assets (%)	6,550	3.4622	7.9426
Revenue growth	6,550	0.0958	0.6702
Loss indicator	6,550	0.1554	0.3623
Asset turnover	6,550	0.7540	0.5084

4. Results

4.1 The pillar decomposition

Table 2 shows the full matrix of Fama-MacBeth estimates for the paper's first central result, with each pillar entered separately against each of the five earnings-quality attributes; inter-pillar correlations reach 0.74, which precludes joint entry without inducing collinearity in the annual cross-sections.

Table 2. Fama-MacBeth cross-sectional estimates by ESG pillar (pillars entered separately)

Dependent variable	β_E	t_E	β_S	t_S	β_G	t_G	T	N
|DA| Modified Jones	−0.000040	−0.85	+0.000116	+1.99*	+0.000072	+1.07	18	6,611
AQ Dechow-Dichev	−0.000168	−4.14***	+0.000009	+0.14	−0.000040	−0.71	10	4,808
Persistence (PER_phi)	+0.000358	+1.16	+0.001100	+7.16***	+0.001082	+3.07***	13	6,169
Predictability (PRED_sigma)	−0.000005	−0.22	−0.000002	−0.06	+0.000039	+2.15*	13	6,169
Smoothness (SMO_NI)	−0.000025	−1.24	−0.000049	−1.74	+0.000065	+2.78**	16	7,137

Note. Each entry reports the Fama-MacBeth (1973) cross-year average of the lagged pillar coefficient across T annual cross-sections (minimum thirty firms per year), with SE = sd(β_t)/√T; */**/*** = 0.10/0.05/0.01. N is the pooled observation count for the Social-pillar specification; the Environmental and Governance samples differ by at most one observation. Controls of Section 3.5 included; year-by-year estimation absorbs time effects by construction.

The matrix supports the differential form of Hypothesis 1. The Social pillar carries the persistence relationship with the highest precision in the matrix (β = +0.001100, t = +7.16), and it carries no smoothness relationship (t = −1.74); a ten-point increment in the Social pillar is associated with a persistence coefficient higher by 0.011, roughly eight per cent of the sample mean of 0.138. The Governance pillar is the only pillar significantly associated with the volatility of scaled net income (β = +0.000065, t = +2.78, p = 0.014), the sign the reduced-masking argument of Hypothesis 1b predicts; the relationship is significant yet modest in magnitude, and the same pillar is also positively and significantly associated with persistence (β = +0.001082, t = +3.07), with a point estimate close to the Social coefficient but less than half its precision. The Environmental pillar is mute on both persistence and smoothness. Yet it is not mute in the matrix: it is the only pillar significantly associated with accrual quality, and in the quality-improving direction (AQ_DD: β = −0.000168, t = −4.14), so that a ten-point Environmental increment corresponds to an accrual-mapping residual lower by 0.0017, roughly three per cent of the sample mean of 0.055. The three relationships together, Social on persistence, Governance on smoothness, Environmental on accrual quality, constitute the differential structure that Hypothesis 1 predicts: no pillar replicates the relationship pattern of another, and no attribute is significantly associated with more than one pillar at conventional levels except persistence.

Two coefficients of the matrix require an honest sentence each. The Social coefficient on discretionary accruals is marginally positive (+0.000116, p = 0.063). The estimand is the unconditional cross-sectional average of the annual ESG-accruals slope, a quantity distinct from the dynamic within-firm effect that the related paper estimates under the system generalised method of moments estimator; the marginal coefficient therefore qualifies the cross-sectional interpretation of the Social pillar rather than contradicting the disciplinary evidence. The Governance coefficient on the predictability attribute is marginally positive

(+0.000039, p = 0.053); because PRED_sigma is a residual volatility, the direction is the same as in the Governance smoothness coefficient, more dispersion rather than less, and the two coefficients are interpreted jointly as one pattern rather than as two findings.

The aggregate score, reported for context, reproduces the structure in diluted form: the Fama-MacBeth coefficient on persistence is +0.000924 (t = 2.90, p = 0.013) and the aggregate smoothness coefficient is insignificant (−0.000037, p = 0.155). Aggregation across pillars averages a precise Social relationship with a mute Environmental one and inverts none of them; the decomposition, not the aggregate, is where the structure resides.

4.2 Within-firm and between-firm triangulation

The triangulation across horizons is the second layer of the substantive structure. Table 3 aligns, for the two attributes that carry Hypothesis 1, the within-firm estimators (two-way fixed effects and the Canay panel quantile regression) against the between-firm estimators (Fama-MacBeth, propensity score matching, entropy balancing).

Table 3. Within-firm versus between-firm estimates: persistence and smoothness

Estimator (identifying variation)	Persistence (PER_phi)	Smoothness (SMO_NI)
Two-way FE (within)	−0.000800 (p = 0.243)	−0.000123 (p = 0.075)*
Canay QR, τ = 0.25 (within)	−0.001472 (p < 0.01)***	−0.000122 (p < 0.01)***
Canay QR, τ = 0.50 (within)	−0.000969 (p = 0.011)**	−0.000102 (p < 0.01)***
Canay QR, τ = 0.75 (within)	−0.000132 (p = 0.754)	−0.000093 (p < 0.01)***
Fama-MacBeth (between)	+0.000924 (p = 0.013)**	−0.000037 (p = 0.155)
PSM ATT, coverage margin (between)	+0.068 (p < 0.01)***	+0.043 (p = 0.389)
Entropy balancing ATT (between)	+0.089 (p < 0.01)***	−0.019 (n.s.)

Note. The regressor is the lagged aggregate ESG Score. FE: firm and year effects, firm-clustered standard errors. Canay (2011): two-step panel quantile estimator, bootstrapped standard errors (500 replications). PSM/EB: treatment is Refinitiv ESG coverage (Section 3.5); ATT in attribute units. Sample sizes per estimator as in Sections 3.1 and 3.5. PSM = propensity score matching; ATT = average treatment effect on the treated.

The persistence column documents a sign inversion across horizons. Within firms, the relationship is negative or null: the fixed-effects coefficient is −0.000800 and insignificant, and the quantile profile is negative and significant in the lower half of the conditional distribution

(−0.001472 at τ = 0.25). Between firms, the relationship is positive and significant under all three cross-sectional devices, with matched coverage differentials of +0.068 and +0.089, one half to two thirds of the sample mean persistence of 0.138. The inversion is the empirical mark of a stock, not a flow: firms with high ESG standing exhibit systematically more persistent earnings than observably similar firms, yet a marginal within-firm movement in a given year is not accompanied by a persistence gain. The pattern is consistent with the reputational mechanism operating through accumulated, multi-year stakeholder commitments whose output is inherently a between-firm quantity.

The smoothness column carries the precision clause of Hypothesis 1. The within-firm relationship is negative, modest, and distributionally homogeneous: the fixed-effects estimate is −0.000123 (p = 0.075) and the Canay profile spans −0.000093 to −0.000122 with significance at the one per cent level at every quantile. This within-firm pattern is attributable to the reputational mechanism, the earnings stability that stable stakeholder relations generate, not to the Governance pillar, which anchors the cross-sectional smoothness coefficient of Table 2, a between-firm quantity of the opposite sign. The two coefficients coexist without contradiction because they answer different questions on different horizons: within a firm, rising ESG standing accompanies slightly calmer earnings; across firms, better-governed firms report the volatility that their operations generate rather than smoothing it away (Leuz et al., 2003; Jin & Myers, 2006). Conflating the two horizons would misattribute the within-firm pattern to governance monitoring, an error the decomposition is designed to prevent; the informative-smoothing counterpoint, under which volatility reduction can impound private information into reported earnings, applies, if anywhere, to the within-firm margin. Under that reading, managers of firms with rising ESG standing would use accrual discretion to signal the permanent component of earnings rather than to conceal its variation, and the modest within-firm coefficient would measure communication rather than concealment; Tucker and Zarowin

(2006) show that smoothing of this kind makes current prices more informative about future earnings. The design here cannot separate informative smoothing from the real earnings stability that stable stakeholder relations generate, and does not need to: both readings place the within-firm pattern outside the governance mechanism, so the precision clause survives under either, a question that recent United Kingdom evidence keeps open (Altin et al., 2026).

4.3 Greenwashing: stable relative prevalence and the measurement-distortion contrast

Table 4 opens the paper's second central result, the greenwashing typology, and reports the classification for the 7,183 firm-years with both classification margins available. The section pairs this result with the paper's measurement contribution: the divergence between the fixed and year-specific classifications, reported below, isolates the distortion that score inflation induces under relative scoring.

Table 4. The greenwashing typology

Panel A: First-tier classification (year-specific medians)

Group	N	%	Mean ESG	Mean |DA|
Genuine (high ESG, low EM)	1,963	27.3	72.9	0.018
Greenwasher (high ESG, high EM)	1,625	22.6	71.8	0.090
Low profile (low ESG, low EM)	1,632	22.7	40.3	0.018
Low ESG, high EM	1,963	27.3	39.4	0.119

Panel B: Refined tier with controversies (N = 7,182)

Group	N	%	Mean ESG	Mean Controversies
Genuine	1,334	18.6	71.2	100
Ambiguous (high ESG, high EM, no controversies)	1,251	17.4	70.8	100
Greenwasher, controversy-refined core	374	5.2	75.0	51.7
Low profile	1,631	22.7	40.3	97.0
Low ESG, high EM	1,963	27.3	39.4	97.1
Other (high ESG, low EM, active controversies)	629	8.8	76.7	53.4

Note. Classification against year-specific medians of the ESG Score and AEM_MJones; the controversies threshold is pooled (Section 3.4). The refined core combines a high declaration, high accruals, and an active controversies signal.

The greenwasher group holds 22.6 per cent of the classified firm-years, and the controversy-refined core, in which an external verification signal contradicts the declaration, holds 5.2 per cent. The core exhibits the profile the construct predicts: the highest mean ESG Score of any group (75.0) combined with a controversies score of 51.7 against a sample mass at 100. Sectorally, greenwashers concentrate in Utilities (30.2 per cent), Energy (28.8 per cent), and Healthcare (25.6 per cent), and are scarcest among Financials (8.8 per cent); the greenwasher indicator is associated with discretionary accruals higher by 0.051 ($p < 0.001$, $N = 7{,}145$) in the classified sample, which confirms that the group isolates economically material accounting behaviour rather than a labelling exercise.

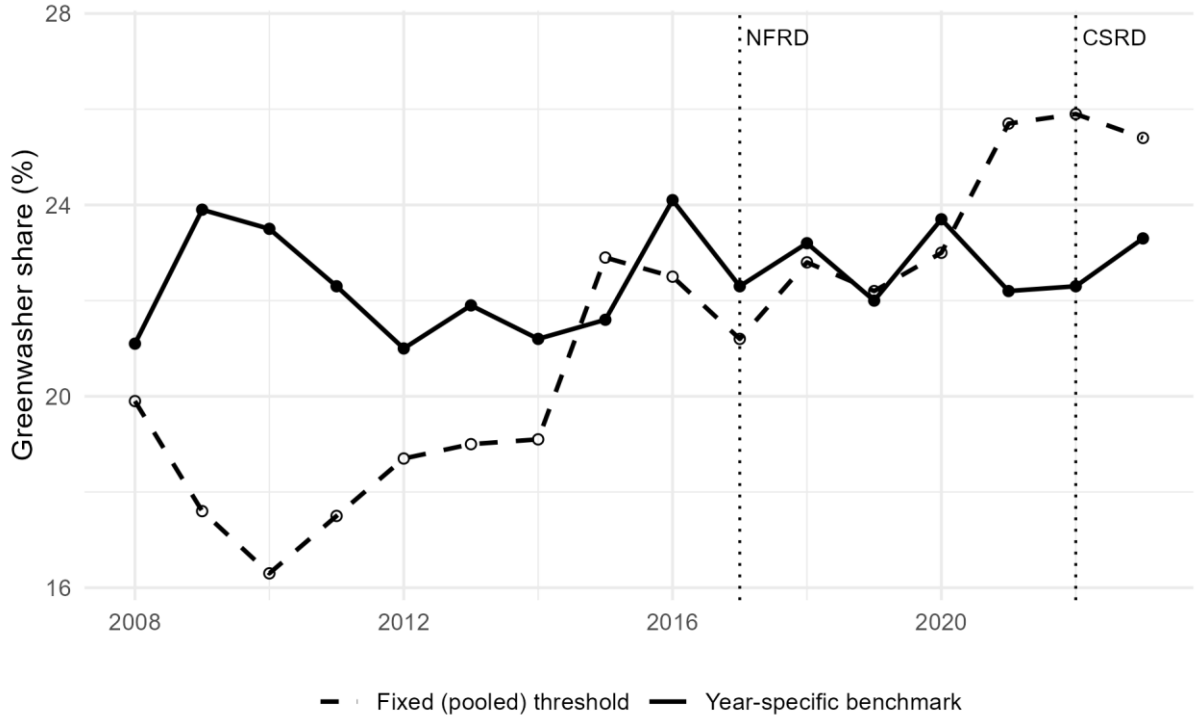
Stability (H2a). Table 5 reports the annual series under both benchmarks, with Wilson intervals for the year-specific series.

Table 5. Annual greenwasher share, 2008–2023: year-specific benchmark versus fixed pooled threshold

Year	N	Share, annual benchmark (%)	Wilson 95% confidence interval	Share, pooled threshold (%)	Difference (pp)
2008	171	21.1	[15.6; 27.8]	19.9	−1.2
2009	176	23.9	[18.2; 30.7]	17.6	−6.3
2010	196	23.5	[18.1; 29.9]	16.3	−7.2
2011	206	22.3	[17.2; 28.5]	17.5	−4.8
2012	214	21.0	[16.1; 27.0]	18.7	−2.3
2013	274	21.9	[17.4; 27.2]	19.0	−2.9
2014	288	21.2	[16.9; 26.3]	19.1	−2.1
2015	301	21.6	[17.3; 26.6]	22.9	+1.3
2016	311	24.1	[19.7; 29.2]	22.5	−1.6
2017	373	22.3	[18.3; 26.7]	21.2	−1.1
2018	569	23.2	[19.9; 26.8]	22.8	−0.4
2019	618	22.0	[18.9; 25.4]	22.2	+0.2
2020	781	23.7	[20.8; 26.8]	23.0	−0.7
2021	809	22.2	[19.5; 25.2]	25.7	+3.5
2022	808	22.3	[19.5; 25.3]	25.9	+3.6
2023	778	23.3	[20.4; 26.4]	25.4	+2.1

Note. N is the annual classified cross-section ($\Sigma = 6{,}873$). Wilson 95 per cent intervals on the year-specific share. The pooled-threshold series applies full-sample medians to every year. Figure 1 renders the two series with the NFRD and CSRD dates marked.

Figure 1. Annual greenwasher share, 2008–2023: year-specific benchmark versus fixed pooled threshold



The year-specific series is flat according to every pre-specified test. Trend-and-regime regression attaches to the trend a coefficient of +0.020 percentage points per year (p = 0.870), to the NFRD indicator +0.26 percentage points (p = 0.791), and to the CSRD indicator +0.02 percentage points (p = 0.966); the weighted variant reproduces the pattern (Table 6). The Chow statistic at the known dates is 0.04 at 2017 (p = 0.959) and, in the predictive form imposed by the two-year post-window, 0.19 at 2022 (p = 0.833); the supplementary Bai-Perron search detects no endogenous break. Yet the stability claim does not rest on non-rejection. The sixteen annual Wilson intervals each contain the sample mean of 22.5 per cent, the band spans 21.0 to 24.1 per cent with a standard deviation of 1.02 percentage points, and the regime point estimates are at most a quarter of that descriptive dispersion. The series is not merely statistically flat; it is economically flat.

Table 6. Tests of H2a: trend, regime, and known-date breaks

Panel A: Regression of the annual share on trend and regime

Term	OLS, Newey-West (lag 2)	Weighted least squares by annual N
Intercept	22.181*** (0.682)	22.019*** (0.876)
Trend	+0.020 (0.120)	+0.046 (0.157)
NFRD (2017+)	+0.257 (0.951)	+0.144 (1.260)
CSRD (2022+)	+0.022 (0.501)	−0.113 (0.767)

Panel B: Chow tests at the known dates

Break date	Form	F	df	p
2017 (NFRD)	Standard	0.042	(2, 12)	0.959
2022 (CSRD)	Predictive (n_post = 2)	0.186	(2, 12)	0.833

Note. Panel A: dependent variable is the annual year-specific greenwasher share in percentage points, T = 16; standard errors in parentheses. Panel B: Chow tests on the share-on-trend regression; the 2022 test uses the predictive form because the post-window is shorter than the parameter count. The supplementary Bai-Perron procedure (share on a constant, minimum segment four years) selects zero breaks by the Bayesian information criterion.

The pooled-threshold series tells the opposite story, and the divergence is the identification device. Under the fixed threshold the apparent share climbs from 16.3 per cent in 2010 to 25.9 per cent in 2022, a trajectory that would read as a regulatory-era expansion of greenwashing. The gap between the two series averages −3.8 percentage points before 2015 and +0.8 afterwards, the mark of a fixed cut applied to a drifting distribution: early cross-sections are mechanically under-classified and late cross-sections over-classified as the score distribution shifts upwards under a relative scoring methodology. The apparent expansion is a measurement distortion, and the contemporaneous benchmark removes it by construction. Any inference about regulatory regimes drawn from a fixed-threshold classification of relative scores inherits this distortion (Benuzzi et al., 2025; Kathan et al., 2025).

Persistence (H2b). Table 7 reports the transition evidence.

Table 7. Greenwasher persistence: transitions and regressions

Panel A: Year-on-year transition matrix (5,989 consecutive pairs, 830 firms)

From \ To	Not greenwasher in t	Greenwasher in t
Not greenwasher in t−1	0.832 (n = 3,848)	0.168 (n = 778)
Greenwasher in t−1	0.502 (n = 684)	0.498 (n = 679)

Panel B: Pooled logit of greenwasher status (principal regression)

Term	β	SE	p	Odds ratio
Status in t−1	1.592	0.084	<0.001	4.91***
NFRD (2017+)	0.149	0.075	0.046	1.16**
CSRD (2022+)	−0.165	0.068	0.016	0.85**
Intercept	−1.662	0.080	<0.001	0.19

Panel C: Firm fixed-effects linear probability model (robustness)

Term	β	SE	p
Status in t−1	+0.022	0.019	0.236
NFRD (2017+)	+0.098	0.017	<0.001
CSRD (2022+)	+0.023	0.013	0.079

Note. Transitions defined over consecutive firm-year pairs within 2008–2023. Panel B: firm-clustered standard errors; the pooled specification estimates total persistence, the sum of the stable firm-level propensity and state dependence; a firm fixed-effects logit is avoided owing to incidental-parameters bias at short T. Panel C: within-firm estimate; the lagged-status coefficient carries the Nickell (1981) caveat and is read as robustness only.

The transition matrix is the principal result. A firm classified as a greenwasher in year t−1 classifies as one again in year t with probability 0.498, against 0.168 for a firm outside the group, and the pooled odds ratio on lagged status is 4.9 ($p < 0.001$). The exits, moreover, barely leave the declaration stratum: of the 1,363 transitions originating in the greenwasher group, 49.8 per cent remain in it, 43.9 per cent move to the genuine group, and only 6.2 per cent cross the ESG boundary into the low-declaration strata. Greenwashing is, in this sample, a recurrent state within the high-ESG population, not a station on the way out of it.

The comparison between Panels B and C locates the source of the persistence. The pooled odds ratio of 4.9 mixes a stable firm-level propensity with genuine state dependence; the within-firm estimate, which differences out the stable propensity, retains no significant lagged-status coefficient (+0.022, $p = 0.236$). The bulk of the persistence therefore resides in permanent

heterogeneity across firms, precisely the firm-level-trait interpretation that Hypothesis 2b advances: some firms are durably positioned in the group, and the year-to-year dynamics add little on top of that positioning.

The regime coefficients in Table 7 estimate a different quantity from the aggregate-share tests of Table 6 and should not be interpreted as tests of H2a. The logit conditions on lagged status, so its NFRD and CSRD coefficients describe shifts in entry and exit intensities conditional on the previous state, while the fixed-effects linear probability model identifies shifts within firms over a panel whose composition changes markedly across regimes, from 171 firms in 2008 to 778 in 2023. Neither speaks to the unconditional share, the subject of H2a. A balanced-panel replication of the transition regressions is available upon request.

4.4 Robustness

Threshold robustness

Table 8 re-runs the classification and the H2a regression under alternative annual cuts.

Table 8. Threshold robustness of the typology and of the stability result

Variant	N GW	% GW	Series min–max	Mean	S.D. (pp)	p trend	p NFRD	p CSRD
Annual medians (canonical)	1,625	22.6	21.0–24.1	22.5	1.02	0.870	0.791	0.966
Annual terciles (top third × top third)	628	8.7	6.5–11.3	8.7	1.33	0.723	0.719	0.665
Annual quartiles (top quarter × top quarter)	291	4.1	2.3–6.3	4.0	1.01	0.511	0.349	0.370
AEM at the annual 60th percentile	1,222	17.0	14.9–18.3	16.7	1.00	0.174	0.865	0.083
AQ_DD, annual medians (suppl., 2013+)	1,169	22.4	20.9–24.3	22.5	0.96	0.139	0.055	0.128

Note. Each variant reclassifies the first tier under the indicated annual cuts and re-estimates the trend-and-regime regression of Table 6, Panel A (Newey-West, lag 2). The AQ_DD variant replaces the accrual margin with the Dechow-Dichev residual and is supplementary because cash-flow coverage restricts its pre-NFRD window to four years (2013–2016).

Stability survives every cut: terciles or quartiles shrink the group, as they must, yet the resulting series remain flat, with dispersions of 1.0 to 1.3 percentage points and no significant trend or regime term. Across the five variants, the fifteen trend-and-regime coefficients include two

below the ten per cent level, the CSRD indicator under the 60th-percentile accrual cut ($\beta = -0.641$, $p = 0.083$) and the NFRD indicator under the AQ_DD variant ($\beta = -1.295$, $p = 0.055$); both are negative, pointing, if anything, to a marginally lower share, and two marginal coefficients in fifteen tests is consistent with sampling variation under the null, the AQ_DD variant offering limited power given its four pre-NFRD years.

The instrument margin

An instrumental-variables specification, instrumenting the lagged ESG Score with the sector leave-one-out peer mean, is reported in full in the Online Appendix (Table A1). The Wu-Hausman statistic does not reject exogeneity ($p = 0.80$), so the associational estimators remain the primary evidence and the specification is retained as a completeness check. Its precise zero on the accruals dimension is consistent with the horizon logic: the peer instrument moves the component of ESG that co-moves within sectors, while the relationships of interest reside in firm-specific accumulated standing, which peer pressure does not shift; the disciplinary margin itself is the object of the related paper.

Regime displacement on the structure

A two-way fixed-effects difference-in-differences design around the NFRD (treatment: firms above the seventy-fifth size percentile; post: 2017 onwards) returns no significant displacement on any structural dimension: persistence ($\delta = -0.001$, $p = 0.950$), predictability (-0.004, $p = 0.224$), smoothness (-0.347, $p = 0.406$), and accrual quality ($+0.003$, $p = 0.689$). This absence is the estimator-level counterpart of the flat share of Section 4.3: the regime neither dislodges the population nor reshapes the attribute structure of Section 4.1. The staggered-adoption caveats of Callaway and Sant'Anna (2021) are examined in the related paper and do not bear on the null pattern reported here.

In sum, the record on the hypotheses is uniform. Hypothesis 1 is supported in its differential form: no attribute is significantly associated with more than one pillar at conventional levels

except persistence, with H1a supported on both clauses and H1b supported through a coefficient that is significant yet modest. Hypothesis 2a is supported across all pre-specified tests, the divergence between the two series standing as the device for the measurement distortion. Hypothesis 2b is supported in its trait form, with the bulk of the persistence residing in permanent firm-level heterogeneity. The marginal terms of Table 2 are qualifications, not contradictions.

5. Concluding remarks

Whether declared sustainability comes with disciplined financial reporting has become a regulatory question as much as an academic one. This paper examined the issue on listed firms of the nineteen Eurozone countries over 2004–2023, measuring earnings quality as five attributes, decomposing the ESG score into its three pillars across two horizons, and testing whether the share of firm-years pairing a high declared score with aggressive accruals is stable across the two European disclosure regimes.

The answer has two parts. The relationship is structured, not uniform. The Social pillar is associated with more durable earnings across firms, the Governance pillar with reported income that tracks performance instead of masking it, and the Environmental pillar with a cleaner mapping of accruals into cash flows. No attribute is shared by more than one pillar except persistence, and each relationship surfaces at the horizon where the corresponding mechanism operates. At the same time, a symbolic core persists. Between 21 and 24 per cent of firm-years pair a high declared score with aggressive accruals, the share crossed the NFRD and the CSRD without a break, and the same firms tend to occupy the group year after year. Under a fixed historical threshold the proportion appears to increase; the increase is a measurement distortion produced by rising scores under relative scoring, not a change in behaviour.

This stability can be reconciled with Aboud et al. (2024), who report that Directive 2014/95 reduced ESG decoupling among covered firms. The two designs measure different things: their decoupling is a continuous distance between disclosure and rating that can narrow on average, whereas the object here is discrete, the share of firm-years in which a strong declaration coexists with aggressive audited accounting, cut against the medians of each year. An average distance can shrink while the occupancy of the extreme group stands still; each result answers a question the other cannot.

This measurement divergence has important implications for how the evidence from the two studies should be interpreted. The design of Aboud et al. (2024), a difference-in-differences framework with US firms as the control group, identifies an average treatment effect on the continuous decoupling margin. It shows that the directive pushed the mean decoupling score downward. The design here identifies the prevalence of a discrete configuration, a question of the upper tail rather than the centre of the distribution. The regulatory pressure that narrows the average gap between disclosure and rating need not dislodge the firms whose declaration and accounting point in opposite directions.

Adjacent evidence points the same way: firms reporting only under compulsion produce more boilerplate disclosure than voluntary reporters (Breijer et al., 2025), and a mandatory regime can lower the average level of accrual-based earnings management (Cui et al., 2025) while the share of firms combining strong scores with aggressive accruals stays where it was.

The prevalence itself should be read against its measurement: an incident-based indicator flags 6.98 per cent of MSCI World firms as greenwashers at least once over 2008–2020, a figure its authors present as a lower bound (Keresztúri et al., 2025), whereas the 21 to 24 per cent here counts firm-years under a two-sided accounting classification. The two figures measure

different units on different margins; their distance shows that measured prevalence depends on the classification device, precisely what the fixed-versus-year-specific contrast formalises.

In short, disclosure mandates move averages more readily than configurations. While the average distance between disclosure and rating may narrow as firms adjust their reporting to align with external benchmarks, the proportion of firms pairing high declared ESG with aggressive accounting may remain unchanged. This asymmetry has direct implications for policy evaluation: in order to assess the effectiveness of disclosure regulation, it is necessary to consider both the intensive margin (i.e. how strongly declared sustainability disciplines reported earnings) and the extensive margin (i.e. who pairs declaration with aggressive accounting). The evidence here suggests that the former is more responsive to regulatory pressure than the latter.

A related paper completes the regulatory picture: Fragoso and Pinheiro (2026) report that regulation compresses the disciplinary margin of the ESG relationship with accruals-based earnings management, while the evidence here indicates that the same pressure leaves the symbolic population in place.

The results support theories that treat sustainability engagement as a binding commitment to stakeholders rather than a decoration, and they qualify the symbolic account without dismissing it: both operate, on different margins. In sum, the paper makes theoretical, methodological, and empirical contributions. Theoretically, it transforms the substantive versus symbolic tension into an observable pattern across pillars, attributes, and horizons. Methodologically, it introduces the fixed-versus-year-specific contrast as a device for the distortion that relative scoring generates under rising scores. Empirically, to the best of our knowledge, it provides the first joint map of the three pillars onto five attributes over two horizons in the Eurozone,

together with formal evidence that greenwashing prevalence is stable and persistent at firm level.

The practical implications follow the two results. Managers should read the pillar map as a guide to how reporting credibility is built: engagement with employees, communities, and customers is associated with more durable earnings; governance quality with reported income that outsiders can trust; environmental engagement with a cleaner accrual process. None of this is visible in the composite score, so boards and audit committees that monitor sustainability through the aggregate number alone discard the attribute-specific signal that the pillars carry.

Investors and analysts should benchmark ESG scores against the distribution of the current year, never against a historical cut-off, because under relative scoring a fixed benchmark manufactures a trend where none exists. The screening recipe the results support is configurational: a strong headline score paired with high discretionary accruals warrants scrutiny of the accrual detail before the score is taken at face value, with particular attention to the sectors where the configuration concentrates, and a firm's past greenwasher status is informative about its current one, since occupancy of the group is persistent rather than transitory.

Authorities enforcing the CSRD should draw the same two lessons. Surveillance thresholds ought to be year-specific, or enforcement will chase a measurement distortion; and because mandates move averages more readily than configurations, enforcement aimed at the persistent core, the firms that occupy the greenwasher group year after year, is better targeted than enforcement aimed at the average reporter. External verification is the margin that works: monitoring by visiting investors curbs greenwashing (Chen et al., 2026), and audited accounting data supply the outsider's counterpart to that verification.

The results are subject to four limitations, each pointing to an extension. First, the scores come from a single provider, and ratings diverge across providers (Berg et al., 2022); replicating the map and the typology with alternative vendors is the natural first step. Second, the sample covers one regulatory space, the Eurozone, chosen for coherence rather than breadth. Third, the greenwashing measure is configurational and peer-relative, faithful to the idea that a firm presents itself as above the prevailing standard while silent on key figures. Fourth, the post-CSRD window closes at two years, so the invariance should be revisited as that regime matures. Future research could also port the year-specific versus fixed-threshold contrast to other relative scoring systems, extend the horizon logic to further attributes, and investigate whether persistent greenwashing predicts outcomes, from financing costs to enforcement, that would render it costly to sustain.

References

Aboud, A., Saleh, A., & Eliwa, Y. (2024). Does mandating ESG reporting reduce ESG decoupling? Evidence from the European Union's Directive 2014/95. *Business Strategy and the Environment, 33*(2), 1305–1320. https://doi.org/10.1002/bse.3543

Altin, M., Al Ani, M. K., Hussainey, K., & Streimikiene, D. (2026). The impact of sustainability expenditures on ESG performance: The moderating role of financial reporting quality. *Corporate Social Responsibility and Environmental Management*. Advance online publication. https://doi.org/10.1002/csr.70588

Ashforth, B. E., & Gibbs, B. W. (1990). The double-edge of organizational legitimation. *Organization Science, 1*(2), 177–194. https://doi.org/10.1287/orsc.1.2.177

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica, 66*(1), 47–78. https://doi.org/10.2307/2998540

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies, 81*(2), 608–650. https://doi.org/10.1093/restud/rdt044

Benuzzi, M., Bax, K., Paterlini, S., & Taufer, E. (2025). Chasing ESG performance: How methodologies shape outcomes. *International Review of Financial Analysis, 104*, Article 104239. https://doi.org/10.1016/j.irfa.2025.104239

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance, 26*(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Breijer, R., Erkens, M. H. R., Orij, R. P., & Vergoossen, R. G. A. (2025). Mandatory versus voluntary non-financial reporting: Reporting practices and economic consequences. *Accounting Forum, 49*(2), 303–335. https://doi.org/10.1080/01559982.2024.2326334

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics, 225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Canay, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal, 14*(3), 368–386. https://doi.org/10.1111/j.1368-423X.2011.00349.x

Carcello, J. V., Hollingsworth, C. W., & Neal, T. L. (2006). Audit committee financial experts: A closer examination using firm designations. *Accounting Horizons, 20*(4), 351–373. https://doi.org/10.2308/acch.2006.20.4.351

Chen, H., & Wang, Z. (2025). Does ESG rating disagreement affect management tone manipulation? International Review of Financial Analysis, 101, Article 104039. https://doi.org/10.1016/j.irfa.2025.104039

Chen, L., Fang, H., Xiong, J., & Yang, Z. (2026). Institutional investors' site visits and corporate greenwashing: Evidence from China. Journal of International Financial Management & Accounting, 37(1), 38–62. https://doi.org/10.1111/jifm.12237

Chih, H.-L., Shen, C.-H., & Kang, F.-C. (2008). Corporate social responsibility, investor protection, and earnings management: Some international evidence. *Journal of Business Ethics, 79*(1–2), 179–198. https://doi.org/10.1007/s10551-007-9383-7

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica, 28*(3), 591–605. https://doi.org/10.2307/1910133

Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review, 97*(1), 147–175. https://doi.org/10.2308/TAR-2019-0506

Clare, A., Pinheiro, C. M., & Pozzolo, A. F. (2025). Does the attentiveness of NEDs to ESG issues influence corporate sustainability? *Corporate Social Responsibility and Environmental Management*. Advance online publication. https://doi.org/10.1002/csr.70118

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management, 37*(1), 39–67. https://doi.org/10.1177/0149206310388419

Cui, X., Li, R., Xue, S., & Zhang, X. (2025). Mandatory versus voluntary: The real effect of ESG disclosures on corporate earnings management. *Journal of International Money and Finance, 154*, Article 103323. https://doi.org/10.1016/j.jimonfin.2025.103323

Dechow, P. M., & Dichev, I. D. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review, 77*(s-1), 35–59. https://doi.org/10.2308/accr.2002.77.s-1.35

Dechow, P. M., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics, 50*(2–3), 344–401. https://doi.org/10.1016/j.jacceco.2010.09.001

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *The Accounting Review, 70*(2), 193–225. https://doi.org/10.2308/TAR-9505096112

Deng, Y., Ong, T. S., & Senik, R. (2024). Trick or treat? A bibliometric literature review of corporate social responsibility and earnings management. *Corporate Social Responsibility and Environmental Management, 31*(5), 4361–4383. https://doi.org/10.1002/csr.2806

Desai, R., & Vyas, J. (2026). Unveiling the past, present, and future of statutory sustainability reporting: An integrated application of bibliometric and qualitative content analysis.

Business Strategy and the Environment. Advance online publication. https://doi.org/10.1002/bse.70935

Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review, 20*(1), 65–91. https://doi.org/10.2307/258887

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science, 60*(11), 2835–2857. https://doi.org/10.1287/mnsc.2014.1984

Eckberg, J., Dorfleitner, G., Kathan, M. C., & Utz, S. (2026). Determinants and forecasting of corporate greenwashing behavior. *Journal of Economic Behavior & Organization, 241*, Article 107354. https://doi.org/10.1016/j.jebo.2025.107354

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*(3), 607–636. https://doi.org/10.1086/260061

Fiechter, P., Hitz, J.-M., & Lehmann, N. (2022). Real effects of a widespread CSR reporting mandate: Evidence from the European Union's CSR Directive. *Journal of Accounting Research, 60*(4), 1499–1549. https://doi.org/10.1111/1475-679X.12424

Fragoso, J., & Pinheiro, C. M. (2026). *ESG performance and earnings quality: Evidence from 20 years of Eurozone listed firms* [Manuscript submitted for publication].

Francis, J., LaFond, R., Olsson, P. M., & Schipper, K. (2004). Costs of equity and earnings attributes. *The Accounting Review, 79*(4), 967–1010. https://doi.org/10.2308/accr.2004.79.4.967

Gao, L., & Yin, H. (2025). Managerial myopia and greenwashing: A machine learning and text analysis approach. International Review of Financial Analysis, 105, Article 104452. https://doi.org/10.1016/j.irfa.2025.104452

Gu, V. X., & Matisoff, D. C. (2026). Symbolic or substantive action: Intent, effort, and results. *Business Strategy and the Environment, 35*(5), 7484–7498. https://doi.org/10.1002/bse.70545

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis, 20*(1), 25–46. https://doi.org/10.1093/pan/mpr025

Hong, Y., & Andersen, M. L. (2011). The relationship between corporate social responsibility and earnings management: An exploratory study. *Journal of Business Ethics, 104*(4), 461–471. https://doi.org/10.1007/s10551-011-0921-y

Huang, Z., Shi, Y., & Jia, M. (2025). Greenwashing: A systematic literature review. *Accounting & Finance, 65*(1), 857–882. https://doi.org/10.1111/acfi.13352

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305–360. https://doi.org/10.1016/0304-405X(76)90026-X

Jin, L., & Myers, S. C. (2006). R² around the world: New theory and new tests. *Journal of Financial Economics, 79*(2), 257–292. https://doi.org/10.1016/j.jfineco.2004.11.003

Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research, 29*(2), 193–228. https://doi.org/10.2307/2491047

Karpoff, J. M., Lee, D. S., & Martin, G. S. (2008). The cost to firms of cooking the books. *Journal of Financial and Quantitative Analysis, 43*(3), 581–611. https://doi.org/10.1017/S0022109000004221

Kathan, M. C., Utz, S., Dorfleitner, G., Eckberg, J., & Chmel, L. (2025). What you see is not what you get: ESG scores and greenwashing risk. *Finance Research Letters, 74*, Article 106710. https://doi.org/10.1016/j.frl.2024.106710

Keresztúri, J. L., Berlinger, E., & Lublóy, Á. (2025). Environmental policy and stakeholder engagement: Incident-based, cross-country analysis of firm-level greenwashing practices. *Corporate Social Responsibility and Environmental Management, 32*(1), 192–211. https://doi.org/10.1002/csr.2945

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review, 91*(6), 1697–1724. https://doi.org/10.2308/accr-51383

Kim, Y., Park, M. S., & Wier, B. (2012). Is earnings quality associated with corporate social responsibility? *The Accounting Review, 87*(3), 761–796. https://doi.org/10.2308/accr-10209

Klein, A. (2002). Audit committee, board of director characteristics, and earnings management. *Journal of Accounting and Economics, 33*(3), 375–400. https://doi.org/10.1016/S0165-4101(02)00059-9

Klein, B., & Leffler, K. B. (1981). The role of market forces in assuring contractual performance. *Journal of Political Economy, 89*(4), 615–641. https://doi.org/10.1086/260996

Krueger, P., Sautner, Z., Tang, D. Y., & Zhong, R. (2024). The effects of mandatory ESG disclosure around the world. *Journal of Accounting Research, 62*(5), 1795–1847. https://doi.org/10.1111/1475-679X.12548

Leggerini, C., Percuoco, M., Ricciardi, I., & Prisco, A. (2026). Circular economy disclosure and ESG performance: Signal, substance, or symbol? *Business Strategy and the Environment, 35*(3), 3672–3690. https://doi.org/10.1002/bse.70334

Leuz, C., Nanda, D., & Wysocki, P. D. (2003). Earnings management and investor protection: An international comparison. *Journal of Financial Economics, 69*(3), 505–527. https://doi.org/10.1016/S0304-405X(03)00121-1

Liu, J., Yu, X., Hu, J., & Ma, J. (2025). The impact of ESG performance on earnings management from a long-term perspective: Evidence from China. *Corporate Social Responsibility and Environmental Management, 32*(6), 7510–7527. https://doi.org/10.1002/csr.70094

Lublóy, Á., Keresztúri, J. L., & Berlinger, E. (2025). Quantifying firm-level greenwashing: A systematic literature review. *Journal of Environmental Management, 373*, Article 123399. https://doi.org/10.1016/j.jenvman.2024.123399

Luo, Y., Zhang, X., & Li, R. (2026). Does ESG report tone influence ESG rating divergence? Evidence from China. Journal of International Financial Management & Accounting, 37(1), 122–150. https://doi.org/10.1111/jifm.12245

Mitchell, R. K., Agle, B. R., & Wood, D. J. (1997). Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *Academy of Management Review, 22*(4), 853–886. https://doi.org/10.2307/259247

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*(3), 703–708. https://doi.org/10.2307/1913610

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica, 49*(6), 1417–1426. https://doi.org/10.2307/1911408

Palas, R., Gafni, D., Solomon, D., & Baum, I. (2025). ESG ratings and financial reporting quality: Why social performance matters. *Finance Research Letters, 86*, Article 108613. https://doi.org/10.1016/j.frl.2025.108613

Persakis, A., Fassas, A. P., & Philippas, D. (2026). Earnings quality and ESG performance in energy and utilities: What really matters? *International Journal of Finance & Economics, 31*(2), 1627–1659. https://doi.org/10.1002/ijfe.70007

Prior, D., Surroca, J., & Tribó, J. A. (2008). Are socially responsible managers really ethical? Exploring the relationship between earnings management and corporate social responsibility. *Corporate Governance: An International Review, 16*(3), 160–177. https://doi.org/10.1111/j.1467-8683.2008.00678.x

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Ștefănescu, C. A., & Suciu, A. L. (2026). Symbolic versus substantive ESG practices: A systematic review and integrative framework. *Corporate Social Responsibility and Environmental Management*. Advance online publication. https://doi.org/10.1002/csr.70722

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review, 20*(3), 571–610. https://doi.org/10.2307/258788

Teti, E., Etro, L. L., & Pausini, L. (2024). Does greenwashing affect company's stock price? Evidence from Europe. *International Review of Financial Analysis, 93*, Article 103195. https://doi.org/10.1016/j.irfa.2024.103195

Tucker, J. W., & Zarowin, P. A. (2006). Does income smoothing improve earnings informativeness? *The Accounting Review, 81*(1), 251–270. https://doi.org/10.2308/accr.2006.81.1.251

Yu, E. P., Luu, B. V., & Chen, C. H. (2020). Greenwashing in environmental, social and governance disclosures. *Research in International Business and Finance, 52*, Article 101192. https://doi.org/10.1016/j.ribaf.2020.101192

Zatini, G., Torricella, N., & della Porta, A. (2026). Costly signals and cheap talk: Measuring the decoupling of ESG routinisation and greenwashing risk. *Business Strategy and the Environment*. Advance online publication. https://doi.org/10.1002/bse.71079

Zharfpeykan, R., & Harrison, J. (2026). Beyond volume: A legitimacy-theory analysis of the substance of corporate sustainability reporting. *Accounting & Finance*. Advance online publication. https://doi.org/10.1111/acfi.70216

Online Appendix

Table A1. Instrumental-variables estimates of the ESG–accruals relationship

	2SLS FE (one instrument)	2SLS FE (two instruments)
Second stage		
ESG Score (instrumented, t−1)	+0.00003	+0.00022
	(0.001)	(0.001)
Firm size (ln assets)	−0.017	−0.018
	(0.012)	(0.012)
Leverage	0.043**	0.043**
	(0.021)	(0.021)
Return on assets	0.000	0.000
	(0.000)	(0.000)
Revenue growth	0.012**	0.012**
	(0.006)	(0.006)
Loss (0/1)	0.007	0.007
	(0.006)	(0.006)
First stage		
Sector leave-one-out peer mean (t−1)		0.402***
		(0.104)
Country-year leave-one-out mean (t−1)		−0.032
		(0.076)
First-stage F	89.9	45.8
Wu-Hausman (p)	0.801	0.679
Sargan (p)		0.103
Firm and year fixed effects	Yes	Yes
N	7,145	7,140
Adjusted R^2	0.209	0.209
Within R^2	0.024	0.024

Note. The table reports in full the two-stage least squares specifications summarised in Section 4.4. The dependent variable is the discretionary-accruals attribute (modified Jones). The instrumented regressor is the one-year-lagged ESG Score; the excluded instruments are the sector leave-one-out peer mean and, in the two-instrument specification, the country-year leave-one-out mean. First-stage coefficients are reported for the two-instrument specification; the single-instrument first stage is summarised by its F statistic. Standard errors clustered at the firm level in parentheses. Wu-Hausman tests the null that the lagged ESG Score is exogenous; its non-rejection motivates the retention of the associational estimators in the main text. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2. Principal estimates excluding financial firms

	Environmental	Social	Governance	Composite ESG (within-firm FE)
Discretionary accruals	−0.000054	+0.000116*	+0.000078	−0.000077
	(−1.11)	(1.89)	(1.13)	(0.000131)
Accrual quality	−0.000184***	+0.000013	−0.000025	−0.000130
	(−4.10)	(0.20)	(−0.44)	(0.000134)
Persistence	−0.000213	+0.000341	+0.001011**	−0.000857
	(−0.60)	(1.63)	(2.78)	(0.001011)
Predictability	−0.000001	+0.000007	+0.000061***	−0.000014
	(−0.04)	(0.20)	(3.24)	(0.000063)
Smoothness	−0.000013	−0.000040	+0.000087***	−0.000101
	(−0.57)	(−1.14)	(3.17)	(0.000068)

Note. The table re-estimates the principal specifications on the sub-panel that excludes financial firms. Columns one to three report Fama-MacBeth coefficients on the lagged pillar scores with t statistics in parentheses (T = 18, 10, 13, 13, and 16 annual cross-sections and N = 6,415; 4,629; 5,739; 5,739; and 6,633 firm-years for the accrual, accrual-quality, persistence, predictability, and smoothness rows). The final column reports the within-firm two-way fixed-effects coefficient on the lagged composite score with firm-clustered standard errors in parentheses (N = 6,330; 4,554; 5,651; 5,651; and 6,552). The Governance and Environmental cells match the full-sample estimates in sign, magnitude, and significance, and the within-firm estimates match in sign and magnitude; the within-firm smoothness coefficient stands at −0.000101 against −0.000123, with the p-value at 0.139 in the smaller panel. The cross-sectional Social persistence coefficient attenuates from +0.001100 (t = 7.16) to +0.000341 (t = 1.63): financial firms combine below-average social scores with below-average earnings persistence and anchor the lower end of the between-firm gradient, and an interaction specification confirms a steeper gradient within financial firms (interaction coefficient +0.00239, p = 0.041), so the full-sample coefficient partly reflects that sector's contribution. The greenwashing typology, with all thresholds recomputed on the non-financial universe, yields a greenwasher share of 22.7 per cent (1,580 of 6,968 classified firm-years), a controversy-refined core of 5.3 per cent (372), and an annual series spanning 20.6 to 24.2 per cent with a dispersion of 1.00 percentage points, against 22.6 per cent, 5.2 per cent, and 21.0 to 24.1 (1.02) in the full universe. * p < 0.10, ** p < 0.05, *** p < 0.01.